```
seq_documentation_block:
LOCUS         CELK09H9      37881 bp      DNA              INV        22-JAN-1998
DEFINITION    Caenorhabditis elegans cosmid K09H9.
ACCESSION     AF043700
VERSION       AF043700.1  GI:2804458
KEYWORDS      .
SOURCE        Caenorhabditis elegans strain=Bristol N2.
  ORGANISM    Caenorhabditis elegans
              Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
              Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE     1   (bases 1 to 37881)
  AUTHORS     Wilson,R., Ainscough,R., Anderson,K., Baynes,C., Berks,M.,
              Bonfield,J., Burton,J., Connell,M., Copsey,T., Cooper,J.,
              Coulson,A., Craxton,M., Dear,S., Du,Z., Durbin,R., Favello,A.,
              Fulton,L., Gardner,A., Green,P., Hawkins,T., Hillier,L., Jier,M.,
              Johnston,L., Jones,M., Kershaw,J., Kirsten,J., Laister,N.,
              Latreille,P., Lightning,J., Lloyd,C., McMurray,A., Mortimore,B.,
              O'Callaghan,M., Parsons,J., Percy,C., Rifken,L., Roopra,A.,
              Saunders,D., Shownkeen,R., Smaldon,N., Smith,A., Sonnhammer,E.,
              Staden,R., Sulston,J., Thierry-Mieg,J., Thomas,K., Vaudin,M.,
              Vaughan,K., Waterston,R., Watson,A., Weinstock,L.,
              Wilkinson-Sproat,J. and Wohldman,P.
  TITLE       2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
              elegans
  JOURNAL     Nature 368 (6466), 32-38 (1994)
  MEDLINE     94150718
REFERENCE     2   (bases 1 to 37881)
  AUTHORS     Madsen,C., Graves,T. and Blair,T.
  TITLE       The sequence of C. elegans cosmid K09H9
  JOURNAL     Unpublished (1998)
REFERENCE     3   (bases 1 to 37881)
  AUTHORS     Waterston,R.
  TITLE       Direct Submission
  JOURNAL     Submitted (20-JAN-1998) Department of Genetics, Washington
              University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT       Submitted by:
                    Genome Sequencing Center
                    Department of Genetics, Washington University,
                    St. Louis, MO  63110, USA, and
                    Sanger Centre, Hinxton Hall
                    Cambridge CB10 1RQ, England
                    e-mail: rw@nematode.wustl.edu and jes@sanger.ac.uk


              NOTICE:  This sequence may not be the entire insert of this clone.
              It may be shorter because we only sequence overlapping sections
              once, or longer because we provide a small overlap between
              neighboring submissions.

              This sequence was finished as follows unless otherwise noted:
              all regions were double stranded or sequenced with an alternate
              chemistry; an attempt was made to resolve all sequencing problems,
              such as compressions and repeats; all regions were covered by
              sequence from more than one subclone


              NEIGHBORING COSMID INFORMATION:


              The 5' cosmid is C45E1, 200 bp overlap;3' cosmid is Y54E10.  Actual
              start of this cosmid is at base position 197 of CELK09H9; actual
              end is at 37881 of CELK09H9.

              NOTES:

              Coding sequences below are predicted from computer analysis, using
              the program Genefinder(P. Green and L. Hillier, ms in preparation).
FEATURES             Location/Qualifiers
     source          1. .37881
                     /organism="Caenorhabditis elegans"
                     /strain="Bristol N2"
```

```
                        /db_xref="taxon:6239"
                        /chromosome="I"
                        /clone="K09H9"
     gene               complement(11887. .13324)
                        /gene="K09H9.4"
     CDS                complement(join(11887. .12723,13250. .13324))
                        /gene="K09H9.4"
                        /note="similar to reverse transcriptase (Pfam; rvt, score;
                        10.85)"
                        /codon_start=1
                        /evidence=not_experimental
                        /protein_id="AAB97571.1"
                        /db_xref="GI:2804461"
                        /translation="MFAVCFSIFICFFLSSFLFLTLVNESFFKPLLFGIYVNDISSIL
                        PEKVACKQFADDTKLYVSTPKAETENNLQAALDAVVDWTRGAKLTLSQSKTVHVTVGR
                        CRKDFKYHLDGYPIERKTITRDLGFLISEKLDFSEHWKKSINLAKFQLANIFNQYSTS
                        NKKLMILLYKTFIRPRLEYGTVVSSPTKKSDEKAIESVQNAFTRRLYSRIKKRYINPT
                        DKDYKTSMERNELFNLQTLSTRRKIIDRIQVIRMNTGKVDLKTSEFFEQQETHTRSKK
                        KYVWRTGKTKLRRHFFVNRTLSTMKQH"
     gene               15300. .16502
                        /gene="K09H9.3"
     CDS                join(15300. .15558,15910. .16502)
                        /gene="K09H9.3"
                        /note="Similar to cuticular collagen"
                        /codon_start=1
                        /evidence=not_experimental
                        /protein_id="AAB97570.1"
                        /db_xref="GI:2804460"
                        /translation="MWKFVIGSVSTAAFFVSVCTIYFSVSMLDELDSFRLSIRDELED
                        WKEVSDDTWQRLNDMTSRNVPKKTNILKEFVRGKRNVGNDQCNCAEPTKNCPAGPPGE
                        KGSLGNPGQPGPDGVDGDNGVDGDVVIHDMPNPKECIKCPAGPPGPPGPPGPLGPRGD
                        KGPSGPRGALGDQGETGPVGEIGDQGPPGSAGRAGPRGQAGQPGTIAIVGLAGRPGPQ
                        GPLGEPGAQGEPGVDGKDGALGAPGRKAENGRPGKRGKDGVAGVPGTRGKEGEDAGYC
                        TCPPRTA"
     gene               complement(17417. .21259)
                        /gene="K09H9.5"
     CDS                complement(join(17417. .17689,17939. .18115,18335. .18665,
                        19906. .19976,20133. .20192,20243. .20356,21065. .21259))
                        /gene="K09H9.5"
                        /note="contains similarity to S. cerevisiae SSF1
                        (SW:P38789) and SSF2 (SW:Q12153)"
                        /codon_start=1
                        /evidence=not_experimental
                        /protein_id="AAB97572.1"
                        /db_xref="GI:2804462"
                        /translation="MKILTVFLLVAISANCATPCDHQDAVKSMECKPLLGQLAGEMAK
                        HADIPPPDELKEFSLLCEQTLACMEQVKCDLLKNATGLIAKTCTGIKLMSGPFGQCID
                        DLRSTPPSLKKYPCGRFLKTEKGRPGNCQMYQDELECTKHLTSDRFTMVLFFDKVPIK
                        EYFTKLFNFYVDFQAENPKYRCVFGKCHVVVSFYFFQKFLRVKYGALHNFYYVKFSYH
                        KILKCAWNQVYCFDAPCLKFVYFFKFARKKWSLACFHELNASKVLLLLEIFLVIPIYM
                        LFLFPWWLMWIGFHLVLILVTIYALRKKKHRFMWPMVLFTLSQFLIWGLLTVIQLLIA
                        FFDTQSFLNFYSQGHHEEFFEKALVVIVVKLLVFAIGAFLFWRFSVYYATKNYYSDRL
                        EGQVSLTSTVILLQ"
     gene               complement(22612. .25875)
                        /gene="K09H9.6"
     CDS                complement(join(22612. .23071,23125. .23648,23857. .23899,
                        24184. .24309,24361. .24437,25223. .25404,25461. .25875))
                        /gene="K09H9.6"
                        /note="contains similarity to human RNA-binding protein
                        FUS/TLS (SW:Q28009)"
                        /codon_start=1
                        /evidence=not_experimental
                        /protein_id="AAB97575.1"
                        /db_xref="GI:2804465"
                        /translation="MKMKKGKKQRKSGAHNRGNVDALSQKDALYHQEKFVKKLQKQKF
                        LENKEIELARQPHCLVIHRGDVGKYVKGLESDLRNLVEPNTAKNLKILKRNNIKDFIV
                        NGAVLGVTNMMVLTSSDASLQLRMMRFSQGPTLSFKVKQYSLARHVVNCQKRPVATDK
                        LFKSSPLVVMNGFGDGTQKHLSLVQTFIQNMFPSINVDTIQLGNLKRCLIVSYDEETD
                        EIQMRHFAIRVVASGLNKSVKKLMQAEKTMGKNIPNLSTYKDISDYFLNPGQFRIRTK
                        LNFSNINHKIIIKKYSIYLNFPSFSSPGQLSDSEFEGDQQEVELPQDISEGRGCGVGQ
                        KSNVRLHEIGPRLTLELVKIEEGIDEGEVLYHKHNAKTPDELIKLRAHMDKKKQMKKR
```

```
                         REQESEQRVIRRLTIVKEQQDAEEAEVKAIRENAARKQAAATGQVEEVENQKEKDREI
                         AMNRERDLKRANEEWGTSEASKRPRYEDSRGGFRGGFRGRGEDRGGFRGRGGDRGGFR
                         GRDRDGGGFRGRSVDRGGFRGGGGGDRGGFRGRSSDRGGDRGGFRGRSGDRDGGFRGGF
                         GGRGGGGFRGGDRGGFRGRGGGGGFRGGRGGDRGGGFRGGRR"
     misc_feature        26104. .26105
                         /note="SL1 trans-splice site; see EST yk418f6.5"
     gene                26108. .28395
                         /gene="K09H9.2"
     CDS                 join(26108. .26329,26385. .26654,26729. .27064,27249. .27346,
                         27392. .27515,27892. .28395)
                         /gene="K09H9.2"
                         /note="coded for by C. elegans cDNA yk418f6.5; coded for
                         by C. elegans cDNA yk414d10.5; coded for by C. elegans
                         cDNA yk418f6.3"
                         /codon_start=1
                         /protein_id="AAB97569.1"
                         /db_xref="GI:2804459"
                         /translation="MNRFVTRSKPTQPTKRVAPAPASGNEENIVAPEPQKTLEEPSVM
                         EKLQHAIREKDAQFTFDILSKNRPFDQAMNKTMTKMEFSPDYSDPSSEFMLLSVDESV
                         IGAFSEGQSLTIRGDSTDDAVLCTDNATFPMKIIESATTVLLLHKCLEAPDSPTLPNF
                         QVKTIGGKCFSTGELCPQVDVLSIGRLKDMLREQELRWDWKDREEEENLKGYKLRDLL
                         NSVQMSVEEVKIALNDLPVVKFSSGRYRYLSHKYRGEMLGLIVELLDEDSIGDVKLEK
                         YFPEKKLNFSIFRCFFPTKIVKTENFSFSEVLLFFSVFSGKKTEFFNFSVVFFKQKSK
                         KNPEKHENSVFRSFFSISRASLRSNLPENVPDEVIDWFLNSRCEQNPEDPFTYRLPEI
                         NLIRDLTVVILYGTQKMPLQQFRELLHKILPYGVTMNIAQLEGVADISDAPFGKVISY
                         LSPEDLPDTIKERMLYLFEYRKLWSMEQLRPYFKDVYKSKVAFDKYVVQNCDYALSES
                         NEMLYCGVR"
     gene                complement(28869. .30417)
                         /gene="K09H9.7"
     CDS                 complement(join(28869. .28914,28963. .29124,29173. .29257,
                         29731. .29862,29998. .30087,30393. .30417))
                         /gene="K09H9.7"
                         /codon_start=1
                         /evidence=not_experimental
                         /protein_id="AAB97573.1"
                         /db_xref="GI:2804463"
                         /translation="MSTAVFANGNQRFPMNSTADSGTFVAKFNAWIDFIGSKKAFFTN
                         LKILAENGIHESVIRALVLHNAKLPIPKCVATESSVEIYNSATISGIFSAFSVRCCKC
                         SDNVTIPNLYKHMKQHVEKTRYSCSTCSGRFETEEAGLHHLLIHRRRRNVKLFDEWDG
                         EAEKRIDQLLMRSVARKRD"
     gene                30469. .33659
                         /gene="K09H9.1"
     CDS                 join(30469. .30737,30856. .30992,31134. .31420,32381. .32483,
                         32992. .33659)
                         /gene="K09H9.1"
                         /note="similar to C. elegans PHA-1 (GB:X74228)"
                         /codon_start=1
                         /evidence=not_experimental
                         /protein_id="AAB97574.1"
                         /db_xref="GI:2804464"
                         /translation="MFTEHLSSELAEPLSPNFRNCETRGKELRPQRAVNIVSRGTREM
                         SGNDITYPREHRLHLKEFLNSRVFAAQFLQFGKIFKKYTSKREMRQILAFSVDFLRSL
                         LSLVNKCLYFYMERNNFLFFVEIGLSEDRSLASLNSIFEEISSNSQPSHHHTNDSITL
                         PYDATIQKIAEKEKNLSKLLKELSIVDSSLEEEKASFEKTQSALLKEKGKIEKDKLML
                         REEAQKQASMNAEMTTIVDSVFNNEAIVSRVLEKTSKNSISNLDLRVQLDCLGFGAIS
                         PIARFDVLPSTQPQSMYILEEVQINVQFAKTLIDELDKDESLFDNFIANSQRIFPTKR
                         VTVEFPMKYIQLSIGNLEKLIKKCMKMVSKEPQRNTTTIFRMHLQDANFSQSVNRVFN
                         IILCDEIPEIFHSSENCKSDDTSDCWHLETIQPIPEQKANQFYGDDTILTGKCFCLKN
                         LRLNSCINYQLLFEKQFLIQFFKEKTGFMRHFYDLVS"
```

alignment_scores:
```
         Quality:    38.00          Length:      7
           Ratio:    5.429            Gaps:      0
Percent Similarity: 100.000    Percent Identity: 100.000
```

alignment_block:
US-09-403-269-7 x CELK09H9   ..

Align seg 1/1 to: CELK09H9 from: 1 to: 37881

```
    1 PheLeuSerGluGlnHisGly 7
```

```
||||||||||||||||||||||||
10126 TTTCTATCCGAGCAACATGGT 10146
```

```
RESULT    3
DRORCPA/c
LOCUS          DRORCPA       5619 bp     mRNA          INV       17-SEP-1999
DEFINITION     Drosophila melanogaster ring canal proteins (kel) mRNA, complete
               cds.
ACCESSION      L08483
VERSION        L08483.1  GI:577275
KEYWORDS       intercellular bridge; ring canal protein.
SOURCE         fruit fly.
  ORGANISM     Drosophila melanogaster
               Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
               Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
               Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE      1  (bases 1 to 5619)
  AUTHORS      Xue,F. and Cooley,L.
  TITLE        kelch encodes a component of intercellular bridges in Drosophila
               egg chambers
  JOURNAL      Cell 72 (5), 681-693 (1993)
  MEDLINE      93201592
   PUBMED      8453663
REFERENCE      2  (bases 1 to 5619)
  AUTHORS      Robinson,D.N. and Cooley,L.
  TITLE        Examination of the function of two kelch proteins generated by stop
               codon suppression
  JOURNAL      Development 124 (7), 1405-1417 (1997)
  MEDLINE      97236487
   PUBMED      9118811
REFERENCE      3  (bases 1 to 5619)
  AUTHORS      Cooley,L.
  TITLE        Direct Submission
  JOURNAL      Submitted (13-JAN-1993) Department of Genetics, Yale University
               School of Medicine, New Haven, CT 06510, USA
COMMENT        On Nov 28, 1994 this sequence version replaced gi:158234.
FEATURES             Location/Qualifiers
     source          1. .5619
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /dev_stage="0-4 hr embryos"
                     /tissue_lib="Nick Brown"
     source          1. .183
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /tissue_type="0-4 hr embryos"
                     /tissue_lib="Nick Brown"
     misc_feature    1. .183
                     /gene="kel"
                     /note="sequence determined from genomic DNA clone"
     gene            1. .5619
                     /gene="kel"
                     /note="kelch"
     CDS             279. .4712
                     /gene="kel"
                     /note="ORF2; putative stop codon suppression by
                     selenocysteine incorporation"
                     /codon_start=1
                     /transl_except=(pos:2346. .2348,aa:OTHER)
                     /product="ring canal protein"
                     /protein_id="AAA53472.2"
                     /db_xref="GI:5902659"
                     /translation="MIALSALLTKYTIGIMSNLSNGNSNNNQQQQQQQQGQNPQQPA
                     QNEGGAGAEFVAPPPGLGAAVGVAAMQQRNRLLQQQQQQHHHHQNPAAEGSGLERGSC
                     LLRYASQNSLDESSQKHVQRPNGKERGTVGQYSNEQHTARSFDAMNEMRKQKQLCDVI
                     LVADDVEIHAHRMVLASCSPYFYAMFTSFEESRQARITLQSVDARALELLIDYVYTAT
                     VEVNEDNVQVLLTAANLLQLTDVRDACCDFLQTQLDASNCLGIREFADIHACVELLNY
                     AETYIEQHFNEVIQFDEFLNLSHEQVISLIGNDRISVPNEERVYECVIAWLRYDVPMR
                     EQFTSLLMEHVRLPFLSKEYITQRVDKEILLEGNIVCKNLIIEALTYHLLPTETKSAR
                     TVPRKPVGMPKILLVIGGQAPKAIRSVEWYDLREEKWYQAAEMPNRRCRSGLSVLGDK
                     VYAVGGFNGSLRVRTVDVYDPATDQWANCSNMEARRSTLGVAALNGCIYAVGGFDGTT
                     GLSSAEMYDPKTDIWRFIASMSTRRSSVGVGVVHGLLYAVGGYDGFTRQCLSSVERYN
                     PDTDTWVNVAEMSSRRSGAGVGVLNNILYRVGGHDGPMVRRSVEAYDCETNSWRSVAD
                     MSYCRRNAGVVAHDGLLYVVGGDDGTSNLASVEVYCPDSDSWRILPALMTIGRSYAGV
```

```
                              CMIDKPMXMEEQGALARQAASLAIALLDDENSQAEGTMEGAIGGAIYGNLAPAGGAAA
                              AAAPAAPAQAPQPNHPHYENIYAPIGQPSNNNNSGSNSNQAAAIANANAPANAEEIQ
                              QQQQPAPTEPNANNNPQPPTAAAPALSQQQQQQQAQPQQPQRILPMNNYRNDLYDRSA
                              ADGVCSAYDVPRAVRSGLGYRRNFRIDMQNGNRCGSGLRCTPLYTNSRSNCQRQRSFD
                              DTESTDGYNLPYAGAGTMRYENIYEQIRDEPLYRTSAANRVPLYTRLDVLGHGIGRIE
                              RHLSSSCGNIDHYNLGGHYAVLGHSHFGTVGHIRLNANGSGVAAPGVAGTGTCNVPNC
                              QGYMTAAGSTVPVEYANVKVPVKNSASSFFSCLHGENSQSMTNIYKTSGTAARMAGHN
                              SPLTPNVSMERASRSASAGAAGSAAAAVEEHSAADSIPSSSNINANRTTGAIPKVKTA
                              NKPAKESGGSSTAASPILDKTTSTGSGKSVTLAKKTSTAAARSSSSGDTNGNGTLNRI
                              SKSSLQWLLVNKWLPLWIGQGPDCKVIDFNFMFSRDCVSCDTASVASQMSNPYGTPRL
                              SGLPQDMVRFQSSCAGACAAAGAASTIRRDANASARPLHSTLSRLRNGEKRNPNRVAG
                              NYQYEDPSYENVHVQWQNGFEFGRSRDYDPNSTYHQQRPLLQRARSESPTFSNQQRRL
                              QRQGAQAQQQSQQPKPPGSPDPYKNYKLNADNNTFKPKPIAADELEGAVGGAVAEIAL
                              PEVDIEVVDPVSLSDNETETTSSQNNLPSTTNSNNLNEHND"
          CDS                 279. .2348
                              /gene="kel"
                              /note="ORF1"
                              /codon_start=1
                              /product="ring canal protein"
                              /protein_id="AAA53471.1"
                              /db_xref="GI:577276"
                              /translation="MIALSALLTKYTIGIMSNLSNGNSNNNNQQQQQQQQGQNPQQPA
                              QNEGGAGAEFVAPPPGLGAAVGVAAMQQRNRLLQQQQQQHHHHQNPAAEGSGLERGSC
                              LLRYASQNSLDESSQKHVQRPNGKERGTVGQYSNEQHTARSFDAMNEMRKQKQLCDVI
                              LVADDVEIHAHRMVLASCSPYFYAMFTSFEESRQARITLQSVDARALELLIDYVYTAT
                              VEVNEDNVQVLLTAANLLQLTDVRDACCDFLQTQLDASNCLGIREFADIHACVELLNY
                              AETYIEQHFNEVIQFDEFLNLSHEQVISLIGNDRISVPNEERVYECVIAWLRYDVPMR
                              EQFTSLLMEHVRLPFLSKEYITQRVDKEILLEGNIVCKNLIIEALTYHLLPTETKSAR
                              TVPRKPVGMPKILLVIGGQAPKAIRSVEWYDLREEKWYQAAEMPNRRCRSGLSVLGDK
                              VYAVGGFNGSLRVRTVDVYDPATDQWANCSNMEARRSTLGVAALNGCIYAVGGFDGTT
                              GLSSAEMYDPKTDIWRFIASMSTRRSSVGVGVVHGLLYAVGGYDGFTRQCLSSVERYN
                              PDTDTWVNVAEMSSRRSGAGVGVLNNILYRVGGHDGPMVRRSVEAYDCETNSWRSVAD
                              MSYCRRNAGVVAHDGLLYVVGGDDGTSNLASVEVYCPDSDSWRILPALMTIGRSYAGV
                              CMIDKPM"
          terminator         2346. .2348
                              /gene="kel"
                              /evidence=experimental
BASE COUNT       1523 a    1405 c    1394 g    1297 t
ORIGIN
```

```
   Query Match           70.4%;   Score 17.6;  DB 3;   Length 5619;
   Best Local Similarity  68.0%;   Pred. No. 2.7e+02;
   Matches   17;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 ccgaattcaargcnatgytnccnyt 25
          |||||||||||: | || :| || :|
Db   4306 CCGAATTCAAAACCATTCTGCCACT 4282
```